

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2002, 20:57:22 ; Search time 1530 seconds

(without alignments)
273.912 Million cell updates/sec

Title: US-09-432-546-15

Perfect score: 39

Sequence: 1 agagatgacctgtgacctggaatgaccttatt 39

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues 22703874

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estcom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_estl:*
11: gp_estl2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24.4	62.6	562	10	AW934005 EST359848
2	24.4	61.5	395	11	BG003920 MR3-GN018
3	23.8	61.0	282	11	W33571 mc54a10.r1
4	23.8	61.0	434	10	AA794057 vr37f06.r
5	23.8	61.0	492	11	BE863099 UI-M-BH0-
6	23.8	61.0	535	11	BF383483 602045165
7	23.8	61.0	786	11	BG966767 602834644
8	23.8	61.0	963	11	W33574 mc54b10.r1
9	23.8	61.0	978	11	BI408490 602963270
10	23.2	59.5	606	11	BF497811 AT1211.5
11	23.2	59.5	609	11	BF499515 AT14081.5
12	23.2	59.5	1101	13	CNS00008 AL076628 Drosophila

13	22.8	58.5	668	10	AW081473
14	22.4	57.4	130	10	AW372361 PM4-BT034
15	22.4	57.4	137	10	AW372352 PM4-BT034
16	22.4	57.4	208	11	H47668 yp75b01.r1
17	22.4	57.4	245	10	AW373822 QY3-BT053
18	22.4	57.4	248	11	T20037 B300F Heart
19	22.4	57.4	251	10	AW373815 QV3-BT053
20	22.4	57.4	286	10	AW1756174 EESTrea40
21	22.4	57.4	349	10	AA084811 zn02b12.s
22	22.4	57.4	361	11	BF733930 PM2-AN008
23	22.4	57.4	397	10	AA307345 EST178239
24	22.4	57.4	414	11	BE831169 PM4-MT003
25	22.4	57.4	436	10	BE780918 601469551
26	22.4	57.4	468	10	AA459606 zxr89604.r
27	22.4	57.4	472	10	BE047641 t239608.y
28	22.4	57.4	496	10	AW328239 ds01b08.x
29	22.4	57.4	502	10	AA429048 zw33812.r
30	22.4	57.4	516	10	BE369822 601284323
31	22.4	57.4	526	10	BE549061 601070019
32	22.4	57.4	528	13	AO373589 RPC11.1-14
33	22.4	57.4	532	13	AO360666 HS-5034.A
34	22.4	57.4	542	10	AA314100 EST185950
35	22.4	57.4	561	10	AA130902 z014h05.r
36	22.4	57.4	565	13	AA0405797 HS-5081.A
37	22.4	57.4	587	10	AA436939 zv72604.r
38	22.4	57.4	600	10	BE391040 601285652
39	22.4	57.4	601	11	BG395154 602457640
40	22.4	57.4	605	10	AA075814 zm75607.r
41	22.4	57.4	612	10	BE383428 601298386
42	22.4	57.4	613	10	BE255445 601116507
43	22.4	57.4	622	11	BG475162 602491717
44	22.4	57.4	624	11	BG707668 602670559
45	22.4	57.4	630	11	BG759481 602712037

ALIGNMENTS

RESULT 1
AW934005 562 bp mRNA EST 18-MAY-2001
LOCUS EST359848 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF57M14 5', mRNA sequence.
ACCESSION AW934005
VERSION AW934005.1 GI:8109406
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 562)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rongling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
Location/Qualifiers
1..562
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF57M14"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"

```

/lab_host="SORR"
/notes="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; ClEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'Immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT      154 a      89 c      125 g      194 t
ORIGIN
Query Match      62.6%; Score 24.4; DB 10; Length 562;
Best Local Similarity 82.4%; Pred. No. 1.1e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY      5 gatggccttggtggcgcttggaatggccctctat 38
      ||| ||||| ||||| ||||| ||||| |||||
Db      260 GATTTTATTGTGTGCCTTCGAAATAGCCTCTGT 293

RESULT 2
BG003920/c      395 bp      mRNA      EST      24-JAN-2001
LOCUS      BG003920
DEFINITION      MR3-GN0186-201100-011-809 GN0186 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BG003920
VERSION      BG003920.1 GI:12444652
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 395)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negal,M.A., da Silva,M.J., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2002663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-GN0186
201100-011-a09&t3=2000-11-20&t4=1)
Seq primer: puc 18 Forward
High quality sequence start: 5
High quality sequence stop: 395.
Location/Qualifiers
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0186"
/dev_stage="Adult"
/notes="Organ: Placenta;normal; Vector: puc18; Site.1: SmaI
; Site.2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT      93 a      118 c      111 g      73 t
ORIGIN

```

Query Match	Similarity	84.4%	Pred. No. 1.5E+02;	DB 11;	Length 395;
Best Local	Similarity	80.0%	Pred. No. 1.7E+02;		
Matches	27;	Conservative	0;	Mismatches	5;
				Indels	0;
				Gaps	0;

Query Match	Similarity	80.0%	Pred. No. 1.7E+02;	DB 11;	Length 282;
Best Local	Similarity	80.0% <td>Pred. No. 1.7E+02; <td></td> <td></td> </td>	Pred. No. 1.7E+02; <td></td> <td></td>		
Matches	28;	Conservative	0;	Mismatches	7;
				Indels	0;
				Gaps	0;

DB	167	AGGACTGGCGCTTGCTTGCCTGCGAAGAGGCCACT	133
RESULT	4		
LOCUS	AA794057/c		
DEFINITION	AA794057	434 bp	mRNA
ACCESSION	VJ7F06.1	Barstead mouse myotubes	MFLRB5 Mus musculus cDNA clone
VERSION	AA794057	IMAGE:1122851	5' mRNA sequence.
KEYWORDS	AA794057.1	GI:2857012	
SOURCE	EST.		
ORGANISM	house mouse.		
	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 434)		
JOURNAL	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubaque,T.,		
COMMENT	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thelsing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R. The WashU-HMI Mouse EST Project Unpublished (1996) Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of MedicineP 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:612187 Seq primer: -26m13 rev2 ET from Amersham High quality sequence stop: 424. Location/Qualifiers 1..434 /organism="Mus musculus" /strain="C3H" /db_xref="taxon:10090" /clone IMAGE:1122851" /clone_id="Barstead mouse myotubes MFLRB5" /cell_line="C2C12" /lab_host="DH10B" /note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCAGATTCGATGAGTGGGACGCGCCGCTTTTATTATTTTTTTTTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."		
BASE COUNT	86 a	142 c	109 g 97 t
ORIGIN			

Query Match	Best Local Similarity	80.0%;	Score 23.8;	DB 10;	length 434;
Matches 28;	Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0;
OY	1	aggagatgcgccttggtgcgccttggaatgcgcctc	35		
Db	237	AGGACCTGGGCTTGGCTTGCCTGGGAAGGCCACT	203		
RESULT	5				
LOCUS	BE863099	432 bp	mRNA	EST	29-SEP-2000

DEFINITION	UT-M-BH0-ajj-e-09-0-0-01.i1 NTH_BMAP_M.S1 Mus musculus cDNA clone
ACCESSION	UT-M-BH0-ajj-e-09-0-0-01.5', mRNA sequence.
VERSION	BE863099
KEYWORDS	BE863099.1 GI:10382729
SOURCE	EST.
ORGANISM	house mouse.
REFERENCE	Mus musculus
AUTHORS	Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae; Murinae; Mus
TITLE	1 (bases 1 to 492)
JOURNAL	Bonaldi,M.F., Lennon,G. and Soares,M.B.
MEDLINE	Normalization and subtraction: two approaches to facilitate gene
COMMENT	discovery Genome Res. 6 (9), 791-806 (1996) 97044477
	Contact: Chin, H
	National Institute of Mental Health
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
	20892-9643, USA
	Tel: 301 443 1706
	Fax: 301 443 9890
	Email: mestremail.nih.gov
	CDNA Library Preparation: M.B. Soares lab
	CDNA clones distribution:
	Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
	should be noted that Bento Soares is generating a small number of
	additional specialized non-redundant arrays of BMAP cDNAs whose
	availability will be considered under appropriate and limited
	collaborative arrangements
	Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers
SOURCE	1..492

	FEATURES	SOURCE
OY	1	aggagatgccttggtgaccttgaagaatgccttc 35
DB	490	AGGAGCTGGGCTTGTCCTGGGAAGAAGCCACT 456
	Query Match	61.0%; Score 23.8; DB 11;
	Best Local Similarity	80.0%; Pred. No. 1.7e+02;
	Matches 28; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
BASE COUNT	112 a	163 c 118 g 99 t
ORIGIN		Location/Qualifiers 1..492 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UI-M-BH0-aj-e-09-0-UI" /clone_lib="NIH_BMAP_M.S1" /dev_stage="27-32 days" /_lab_host="DHL0B (Life Technologies)" /note="Vector: pYT73-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_M.S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain."

RESULT	6								
LOCUS	BF383483/c								
DEFINITION	BF383483	535 bp	mRNA	EST	27-NOV-2000				
VERSION	602045165F1	NCT_CGAP_L19	Mus musculus	cDNA clone IMAGE:4194352	5'				
KEYWORDS	BF383483								
SOURCE	BF383483.1	GI:11364788							
ORGANISM	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclirognathii; Muridae; Murinae; Mus								

FEATURES	source
<p>AUTHORS TITLE JOURNAL COMMENT</p>	<p>1 (bases 1 to 535) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: http://mgc.nci.nih.gov/. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN1 at: http://image.llnl.gov plate: L1AM1002 row: b column: 17 High quality sequence stop: 533. Location/Qualifiers</p>
<p>BASE COUNT ORIGIN</p>	<p>108 a 182 c 131 g 113 t 1 others</p>
<p>Query Match Best Local Similarity Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0.</p>	<p>61.0%; Score 23.8; DB 11; Length 535; 80.0%; Pred. No. 1.7e+02; 0; Indels 7; Gaps 0.</p>
<p>RESULT 7 LOCUS BG966767/c DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM</p>	<p>1 aggaagatgccttggtgcttggaatgacgctct 35 Db 286 AGGAGCTGGGCTTGCTTGCTTGGAAGAGCCACT 252</p> <p>786 bp mRNA EST 12-JUN-2001 6028346441 NC1-CGAP-Co24 Mus musculus cDNA clone IMAGE:4988986 5', mRNA sequence. BG966767 BG966767.1 GI:14354404 EST. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 786) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: http://mgc.nci.nih.gov/. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (MLN1) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN1 at: http://image.llnl.gov plate: L1AM1002 row: b column: 11 High quality sequence start: 6 High quality sequence stop: 779. Location/Qualifiers</p>
<p>FEATURES source</p>	<p>1. 786 /oranism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone_image="4194352" /clone_id="NC1-CGAP_L19" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: liver; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NC1-CGAP Library."</p>

BASE COUNT	170 a	264 c	195 g	157 t	
ORIGIN					
Query Match	61.0%;	Score 23.8;	DB 11;	Length 786;	
Best Local Similarity	80.0%;	Pred. No. 1.8e+02;			
Matches	28;	Conservative	0;	Mismatches	7;
				Indels	0;
				Gaps	0;
OY	1	aggagatggccttgatggccttggaatggcctc	35		
Db	556	AGGACCTGGGCTTGCTTCCTGGGAAGAGGCACCT	522		
RESULT	8				
W33574/c					
LOCUS	W33574	963 bp	mRNA	EST	13-MAY-1996
DEFINITION	mc4b10.r1 Soares mouse embryo NBME13.5	14.5	Mus musculus	CDNA	
LOCATION	clone IMAGE:552315	5'	mRNA sequence.		
ACCESSION	W33574				
VERSION	W33574.1	GI:1315625			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 963)				
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,				
	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,				
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,				
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and				
	Waterston,R.				
TITLE	The WashU-HMNI Mouse EST Project				
JOURNAL	unpublished (1996)				
COMMENT	Contact: Marra M/Mouse EST Project				
	WashU-HMNI Mouse EST Project				
	Washington University School of Medicine#				
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
	Tel: 314 286 1800				
	Fax: 314 286 1810				
	Email: mouseest@watson.wustl.edu				
	This clone is available royalty-free through LNL; contact the				
	IMAGE Consortium (info@image.llnl.gov) for further information.				
	MGI:224115				
	Seq primer: ETPrimers				
FEATURES	High quality sequence stop: 321.				
source	Location/Qualifiers				
	1..963				
	/organism="Mus musculus"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:352315"				
	/clone_id="Soares mouse embryo NBME13.5 14.5"				
	/sex="unknown"				
	/issue_type="embryo"				
	/dev_stage="13.5-14.5dpc total fetus"				
	/lab_host="DH10B"				
	/note="Vector: pT73D-Pac (Pharmacia) with a modified				
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA				
	was primed with a Not I - oligo(dT) primer 15'				
	TGTCACCAATCTGACATGGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT				
	T 3], on equal amounts of mRNA from 2 13.5dpc and 2				
	14.5dpc embryos [total RNA provided by Minoru Ku, Wayne				
	State Univ., from 2] ; double-stranded cDNA was ligated to				
	Eco RI adaptors (Pharmacia), digested with Not I and				
	cloned into the Not I and Eco RI sites of the modified				
	pT73 vector. Library went through one round of				

BASE COUNT	223 a	224 c	261 g	246 t	9 others
ORIGIN	normalization, and was constructed by Bento Soares and M. Fatima Bonaldi.				

Query Match	61.0%	Score 23.8	DB 11	Length 963
Best Local Similarity	80.0%	Pred. No. 18e+02		
Matches	28	Conservative	0	Mismatches 7
				Indels 0
				Gaps 0
Oy	1	aggagatgccttgcttgagccttggaatgagcctc	35	
Db	167	AGGACCTGGCTTGCTTGCTTGAGGAGGCCACT	133	

RESULT	9
B1408490/c	
LOCUS	
DEFINITION	B1408490 978 bp mRNA
VERSION	60296332.f01 NC1_CGAP_LnJ3 Mus musculus cdna clone IMAGE:5118857 5'
ACCESSION	mRNA sequence.
KEYWORDS	B1408490 B1408490.1 GI:15169413
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 978)	NIH-MGC	http://mgc.nci.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.			

Issue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldi, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM11289 row: k column: 18
 High quality sequence start: 4
 High quality sequence stop: 917.

Location/Qualifiers
SOURCE

1. 978
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="TMASE:5116857"
/clone_11b="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DHA10B (phage-resistant)"
/note="Organ: Lung; Vector: pPT73D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - Oligo(dT) primer [5', TGTTCACCATCTGAAGTGAGCGGCCGCCTCTCTTTTATTTTTTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

183 a 335 c 237 g 223 t

```

Query Match          61.0%; Score 23.8; DB 11; Length 978;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 aggaagatgccttgatgccttgaaatggccttc 35
||||| ||| |||||| |||| |||| |||| ||

```

Db 124 AGGAGCTGGGCTTGTTGCCTGGGAAGAGGCCACT 90

RESULT	10
BF497811/c	
LOCUS	BF497811 606 bp mRNA EST 19-APR-2001
DEFINITION	AT12111.primr AT Drosophila melanogaster adult testes POFB7
	Drosophila melanogaster cDNA clone AT1211 5 similar to CG13472:
	Fnan0013472 located on: 3L 70F5-70F6;: 04/09/2001, mRNA sequence
ACCESSION	BF497811
VERSION	BF497811.2 GI:13689723
KEYWORDS	
SOURCE	EST: fruit fly.

ORGANISM
Urosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 606)

REFERENCE
AUTHORS
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Beraman,
B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, Y., Fafan,
D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nuno, J., Paclebo, J.,
Pargacs, V., Park, S., Phouenavong, S., Wan, K., Yu, C., Lewis, S. E.,
Celisner, S. and Rubin, G. M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)

TITLE
JOURNAL
COMMENT
On Dec 6, 2000 this sequence version replaced gt:11581112.
Contact: Stapleton, M.
BDB1

Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 hit genomic AB003533: arm:3L [14543825..14618409]
 estimated-cyto:70E7-71A2: 04/09/2001
 plate: AT 121 row: A column: 11
 High quality sequence stop: 549.
 Location/Qualifiers
 1..606
 source

```

155 a 155 c 185 g 111 t
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="Ar12111"
/clone_jib="Ar Drosophila melanogaster adult testes potb7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates Ar.10-Ar.120: DH5-alpha. Plates
Ar.121-Ar.319: DH5-alpha TonaA"
/note="Organ: ADULT testes; Vector: potb7; site.1: EcoRI;
site.2: Xho1; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potb7. Plasmid cDNA library."

```

Query Match	59.5%	Score 23.2;	DB 11;	Length 606;
Best Local Similarity	89.3%	Pred. No. 2.8e+02;		
Matches	25;	Conservative	0;	Mismatches 3;
				Indels 0;
				Gaps 0;
QY	8	ggacctgtgagccttggaatgacctc	35	
Db	600	ggccttgatgacacgcgcacatgaccctc	573	

RESULT	11
LOCUS	BF499515/c
DEFINITION	BF499515 609 bp mRNA EST 19-Apr-2001
ACCESSION	AT11081.1 Sprine AT <i>Drosophila</i> melanogaster adult testes protein <i>Drosophila</i> melanogaster cDNA clone AT11081 5 similar to CG13472: FBan0013472 Located on: 3L 70F5-70F6; 04/09/2001, mRNA sequence BF499515

VERSION BF499515.2 GI:13691376
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 609)
 Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman, R., Carlson, J., Chame, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frisoe, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Mira, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S., and Rubin, G. M.
 BDGP/HHMI AT Drosophila EST Project
 Unpublished (2000)
 On Dec 6, 2000 this sequence version replaced gi:11582816.
 COMMENT Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: <http://www.fruitfly.org/EST/estefruitfly.berkeley.edu>
 hit genomic AEO03533: arm:3L [14543825,14818409]
 estimated-cyco:70E7-71A2: 04/09/2001
 Plate: AT.140 Row: G Column: 9
 High quality sequence stop: 581.
 Location/Qualifiers
 1..609
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="AT114081"
 /clone_1lb="AT Drosophila melanogaster adult testes POTB7"
 /sex="male"
 /dev_stage="0-3 day old Ore-R males"
 /lab_host="Plates AT.10-AT.120: DH5-alpha Tona"
 AT.121-AT.319: DH5-alpha Tona"
 /note="Organ: ADULT testes; Vector: POTB7; Site:1: EcoRI; Site:2: XhoI; The mRNA for the testis library was made from testes and seminal vesicles hand dissected from 0-3 day old Ore-R males. RNA kindly provided by the lab of Margaret Fuller. Sized fractionated cDNAs were directly ligated into POTB7. Plasmid cDNA library."
 BASE COUNT 157 a 155 c 186 g 111 t
 ORIGIN
 Query Match 59.5%; Score 23.2; DB 11; Length 609;
 Best Local Similarity 89.3%; Pred. No. 2.8e+02;
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 ggccttgatgacctggaatgacctct 35
 ||||||||| |||||||||
 Db 602 ggccttgatgacctggaatgacctct 575
 RESULT 12
 CENS0008/c 1101 bp DNA GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
 DEFINITION BAC38619 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 ACCESSION AL076628 GI:4956105
 VERSION AL076628
 SOURCE fruit fly.
 KEYWORDS Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 Genoscope.
 TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqlref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammotter in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 FEATURES
 source
 1..1101
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_1lb="RPCI-98"
 /clone="BAC38619"
 /note="end : TET3"
 BASE COUNT 301 a 283 c 299 g 185 t 33 others
 ORIGIN
 Query Match 59.5%; Score 23.2; DB 13; Length 1101;
 Best Local Similarity 89.3%; Pred. No. 2.9e+02;
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 ggccttgatgacctggaatgacctct 35
 ||||||||| |||||||||
 Db 362 ggccttgatgacctggaatgacctct 335
 RESULT 13
 AM081473 668 bp mRNA EST 14-OCT-1999
 LOCUS xc31809.x1 NCL CGAP Col8 Homo sapiens cDNA clone IMAGE:2585896 3'
 DEFINITION similar to gb:U24097 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, CW*1202 ALPHA CHAIN (HUMAN);, mRNA sequence.
 ACCESSION AM081473
 VERSION AM081473.1 GI:6036625
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 668)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@emall.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christina Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNT at:
www.bio.lnl.gov/bdrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 415.
 Location/Qualifiers
 1..668
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2585896"

```

/clone_1lb="NCI_CGAP_Co18"
/lisub_type="moderately differentiated adenocarcinoma"
/lab_host="PH10B"
/Note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.26 kb. Library constructed by Life
Technologies. Normalized versions of this library named
NCI_CGAP_Co19 (cot 50), NCI_CGAP_Co20 (cot 500), and
NCI_CGAP_Co21 (cot >500)."
BASE COUNT      178 a      163 c      160 g      164 t      3 others
ORIGIN

Query Match      58.5%; Score 22.8; DB 10; Length 668;
Best Local Similarity 79.4%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 gatggccttggtggccttggaatgctctctat 38
||||| ||||||||| ||||| |||
Db 620 gatggccttggtggccttggaatgctctctat 653

RESULT 14
AM372361      130 bp      mRNA      EST      04-FEB-2000
LOCUS
DEFINITION PM4-BT0341-251199-002-C12 BT0341 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM372361
VERSION AM372361.1 GI:6877119
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 130)
HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM4&c2=PM4-BT0341-
251199-002-C12&t3=1999-11-25&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 130.
Location/Qualifiers
1. 130
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="BT0341"
/dev_stage="Adult"
/Note="Organ: breast; Vector: puc18; Site: 1: SmaI; Site: 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      18 a      33 c      43 g      36 t
ORIGIN

Query Match      57.4%; Score 22.4; DB 10; Length 130;
Best Local Similarity 81.2%; Pred. No. 4.8e+02;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 gatggccttggtggccttggaatgctctct 36
||||| ||||||||| ||||| |||
Db 96 gatggccttggtggccttggaatgctctct 127

```

```

||||| ||||||||| ||||| |||||
Db 89 gatggccttggtggccttggaatgctctct 120

RESULT 15
AM372352      137 bp      mRNA      EST      04-FEB-2000
LOCUS
DEFINITION PM4-BT0341-251199-002-E10 BT0341 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM372352
VERSION AM372352.1 GI:6877110
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 137)
HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM4&c2=PM4-BT0341-
251199-002-E10&t3=1999-11-25&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 63.
Location/Qualifiers
1. 137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="BT0341"
/dev_stage="Adult"
/Note="Organ: breast; Vector: puc18; Site: 1: SmaI; Site: 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      19 a      35 c      45 g      38 t
ORIGIN

Query Match      57.4%; Score 22.4; DB 10; Length 137;
Best Local Similarity 81.2%; Pred. No. 4.8e+02;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 gatggccttggtggccttggaatgctctct 36
||||| ||||||||| ||||| |||
Db 96 gatggccttggtggccttggaatgctctct 127

```

Search completed: January 29, 2002, 21:50:19
Job time: 3177 sec

